Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Magdalena Hrochová

Leveneův test shodnosti rozptylů

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Ing. Marek Omelka, Ph.D.
Studijní program: Matematika
Studijní obor: Obecná matematika

Praha 2014

Ráda bych poděkovala vedoucímu své bakalářské práce, panu Ing. Marku Omelkovi, Ph. D., za odborné vedení, cenné rady a doporučení vhodné literatury. Dále bych ráda poděkovala těm, kteří mne podporovali celý můj život a díky kterým mohu studovat to, co mne zajímá - svým rodičům. V neposlední řadě bych ráda poděkovala Mgr. Petru Lasákovi za cenné rady a korekturu.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne .... Podpis autora

Název práce: Leveneův test shodnosti rozptylů

Autor: Magdalena Hrochová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Ing. Marek Omelka, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Tato bakalářská práce se zabývá Leveneovým testem shodnosti rozptylů k nezávislých náhodných výběrů a jeho modifikacemi. V úvodu práce je popsána analýza rozptylu (ANOVA - z anglického ANalysis Of VAriance) testující rovnost středních hodnot k nezávislých náhodných výběrů, kterou Leveneův test využívá. Následuje popis vlastního testu a ověření předpokladů pro použití ANOVY. Vzhledem k faktu, že původní Leveneův test není spolehlivý, pokud data pocházejí z některých specifických rozdělení (například χ^2), jsou v práci uvedeny převzaté či nově navržené modifikace umožňující dosažení lepší spolehlivosti. Hlavní část práce je věnována odůvodnění správnosti popsaného testu včetně ověření pomocí simulací. V závěru práce jsou porovnány výsledky simulací pro jednotlivé varianty Leveneova testu a další možné testy shodnosti rozptylů.

Klíčová slova: Levene, k-výběrový problém, shoda rozptylů, simulace

Title: Levene's Test for Equality of Variances

Author: Magdalena Hrochová

Department: Department of Probability and Mathematical Statistics

Supervisor: Ing. Marek Omelka, Ph.D., Department of Probability and Mathematical Statistics

Abstract: In the presented bachelor thesis, we focused on the Levene's test and its modifications that are used to assess the equality of variances for k independent random samples. At the beginning, we described Analysis of variance (ANOVA) that is a method for analyzing differences between group means for k independent random samples. The following part contains the original Levene's test description, including a discussion about the assumptions' verification for using ANOVA. Considering the fact that the original test is not confident in case of samples from specific distributions (e.g. a chi-square distribution) we summarize some known and suggest some new modifications. The main part of the thesis is dedicated to experimental simulations. In the conclusion, we discuss the simulation results for different versions of the Levene's test and other similar tests for equality of variances.

Keywords: Levene, k-sample problem, equality of variances, simulations

Obsah

Úvod **2**

1 ANOVA **3**

2 Leveneův test **4**
2.1 Polo-normální rozdělení 4
2.2 Původní Leveneův test 5
2.3 Předpoklady analýzy rozptylu 6
2.4 Další varianty Leveneova testu 7

3 Další testy shodnosti rozptylů **9**
3.1 Bartlettův test 9
3.2 Hartleyův a Cochranův test 10

4 Simulační experimenty **11**
4.1 Návrh simulačních experimentů 11
4.2 Výsledky simulací 12

Závěr **18**

Seznam použité literatury **19**

Seznam tabulek **20**

Úvod

Tématem této práce je Leveneův test (viz Levene, 1960), který je dnes asi nejpoužívanějším testem shodnosti rozptylů k nezávislých náhodných výběrů. V praxi je třeba testovat shodu rozptylů například pokud máme data z několika laboratoří a chceme vědět, zda všechny laboratoře měří se stejnou přesností.

Jednoduchý test shody rozptylů dvou nezávislých náhodných výběrů pocházejících z normálního rozdělení je F-test, který nyní popíšeme.

Předpokládejme, že máme $X_1,\ldots,X_n$ náhodný výběr z $\mathrm{N}(\mu_1,\sigma_1^2)$, který je nezávislý s náhodným výběrem $Y_1,\ldots,Y_m$ z $\mathrm{N}(\mu_2,\sigma_2^2)$. Nechť navíc $m \geq 2, n \geq 2$ a oba rozptyly jsou kladné. Chceme testovat hypotézu

$$H_0 : \sigma_1^2 = \sigma_2^2$$

proti alternativě, že tomu tak není. Definujeme testovou statistiku

$$F_2 = \frac{S_X^2}{S_Y^2},$$

kde $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2, \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, analogicky pro S_Y^2. Za platnosti nulové hypotézy má veličina F_2 Fisherovo-Snedecorovo rozdělení o $n-1$ a $m-1$ stupních volnosti (viz Anděl, 2003, Věta 8.3).

Mohou však nastat situace, kdy je třeba testovat shodnost více než dvou rozptylů, potom se F-test nedá použít. V takovém případě je nutno zvolit jiný test k tomu vhodný, například právě Leveneův. V literatuře, která se jím zabývá, se nejčastěji dočteme, jak hypotézu shodnosti rozptylů testovat, málokdy je však popsán princip testu. To je rovněž jedním z cílů této práce.

Nejprve nastíníme metodu analýzy rozptylu, poté popíšeme vlastní Leveneův test a některé jeho modifikace. Následně jsou uvedeny některé další testy shodnosti rozptylů. V závěru práce se pomocí simulací pokusíme nahlédnout, jak se jednotlivé testy chovají v závislosti na tom, ze kterého rozdělení pocházejí náhodné výběry.

Kapitola 1

ANOVA

Analýza rozptylu (z anglického ANalysis Of VAriance) je statistická metoda, která ověřuje hypotézu shodnosti středních hodnot k výběrů z normálního rozdělení se stejným rozptylem. Ta se používá i v Leveneově testu shodnosti rozptylů, proto ji zde alespoň stručně popíšeme.

Předpokládejme, že máme k nezávislých náhodných výběrů z normálního rozdělení o rozsazích $n_1, \dots, n_k$. Tedy máme

$$\begin{gathered} Y_{11}, \dots, Y_{1n_1} \text{ výběr z } \mathrm{N}(\mu_1, \sigma^2), \\ \vdots \\ Y_{k1}, \dots, Y_{kn_k} \text{ výběr z } \mathrm{N}(\mu_k, \sigma^2). \end{gathered}$$

Povšimněme si, že předpokládáme stejné rozptyly všech výběrů.

Označíme

$$Y_{i.} = \sum_{j=1}^{n_i} Y_{ij}, \qquad y_{i.} = \frac{1}{n_i} Y_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, \quad i = 1, \dots, k$$

a dále

$$N = n_1 + \ldots + n_k, \qquad Y_{..} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}, \qquad y_{..} = \frac{1}{N} Y_{..}.$$

Testujeme hypotézu

$$H_0 : \mu_1 = \ldots = \mu_k$$

proti alternativě, že tomu tak není. Testová statistika

$$F = \frac{(N-k) \sum_{i=1}^{k} n_i (y_{i.} - y_{..})^2}{(k-1) \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - y_{i.})^2} \tag{1.1}$$

má za platnosti nulové hypotézy Fisherovo-Snedecorovo rozdělení (neboli F rozdělení) o $k-1$ a $N-k$ stupních volnosti (viz Anděl, 2007, Kapitola 10.4). Tedy zamítáme hypotézu H_0 na hladině α, pokud je $F \geq F_{k-1,N-k}(\alpha)$, kde $F_{k-1,N-k}(\alpha)$ je příslušná kritická hodnota[1], kterou lze nalézt například v (Anděl, 2003, Tabulka T6).

[1] Kritickou hodnotou je myšlen $1-\alpha$ kvantil příslušného rozdělení.

Kapitola 2

Leveneův test

V této kapitole nejprve uvedeme některá pomocná tvrzení o absolutní hodnotě veličiny, která má normální rozdělení. Poté popíšeme a vysvětlíme princip původního Leveneova testu. Ten zásadním způsobem využívá analýzu rozptylu (viz Kapitola 1), proto prodiskutujeme, zda splňuje její předpoklady. Na závěr jsou uvedeny některé možné modifikace Leveneova testu.

2.1 Polo-normální rozdělení

Věta 1. *Nechť náhodná veličina X má normální rozdělení s nulovou střední hodnotou a rozptylem $\sigma^2 > 0$. Potom náhodná veličina $Y = |X|$ má střední hodnotu $\sqrt{\frac{2}{\pi}}\,\sigma$ a rozptyl $\left(1 - \frac{2}{\pi}\right)\sigma^2$.*

Důkaz. Hustota náhodné veličiny X je $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}$, potom

$$\begin{aligned}
\mathrm{E}Y = \mathrm{E}|X| &= \int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x = 2\int_0^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x \\
&= 2\sigma^2 \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\frac{x}{\sigma^2}\,\mathrm{d}x = \begin{bmatrix} y = \frac{x^2}{2\sigma^2} \\ \mathrm{d}y = \frac{x}{\sigma^2}\mathrm{d}x \end{bmatrix} \\
&= 2\sigma^2 \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-y}\,\mathrm{d}y \\
&= \frac{2\sigma^2}{\sqrt{2\pi\sigma^2}}\left[-\mathrm{e}^{-y}\right]_{y=0}^{\infty} = \sqrt{\frac{2}{\pi}}\,\sigma.
\end{aligned}$$

Dále

$$\begin{aligned}
\mathrm{var}(Y) &= \mathrm{E}Y^2 - (\mathrm{E}Y)^2 = \mathrm{E}X^2 - \frac{2}{\pi}\sigma^2 = \mathrm{var}(X) + (\mathrm{E}X)^2 - \frac{2}{\pi}\sigma^2 \\
&= \sigma^2 + 0 - \frac{2}{\pi}\sigma^2 = \left(1 - \frac{2}{\pi}\right)\sigma^2.
\end{aligned}$$

Definice 1. *Rozdělení náhodné veličiny Y z Věty 1 nazveme* polo-normální rozdělení s parametrem σ^2.

Máme-li náhodný výběr $X_1,\dots,X_n$ z $\mathrm{N}(\mu,\sigma^2)$, potom je $X_1-\mu,\dots,X_n-\mu$ náhodný výběr z $\mathrm{N}(0,\sigma^2)$. Absolutní hodnoty $|X_1-\mu|,\dots,|X_n-\mu|$ budou pak výběrem z polo-normálního rozdělení s parametrem σ^2.

V praxi většinou hodnotu μ neznáme a proto ji nahrazujeme například nestranným odhadem $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$. Potom ale $X_1-\bar{X},\dots,X_n-\bar{X}$, respektive $|X_1-\bar{X}|,\dots,|X_n-\bar{X}|$ není náhodný výběr, protože odečítáním výběrového průměru $\bar{X}$ vnášíme závislost. V následující části uvidíme, proč jsou tyto znalosti důležité.

2.2 Původní Leveneův test

Předpokládejme, že máme k nezávislých náhodných výběrů z normálního rozdělení o rozsazích $n_1,\dots,n_k$. Tedy máme

$$\begin{aligned} Y_{11},\dots,Y_{1n_1} &\text{ výběr z } \mathrm{N}(\mu_1,\sigma_1^2)\\ &\vdots\\ Y_{k1},\dots,Y_{kn_k} &\text{ výběr z } \mathrm{N}(\mu_k,\sigma_k^2). \end{aligned}$$

Chceme testovat hypotézu

$$H_0:\sigma_1^2=\ldots=\sigma_k^2$$

proti alternativě, že tomu tak není. Uvažujme stejné značení jako v Kapitole 1.

Věta 2. *Definujme náhodné veličiny $Z_{ij}=|Y_{ij}-y_{i.}|$ pro $i=1,\dots,k$, $j=1,\dots,n_i$. Potom Z_{ij} má polo-normální rozdělení s parametrem $\left(1-\frac{1}{n_i}\right)\sigma_i^2$.*

Důkaz. Označme $H_{ij}=Y_{ij}-y_{i.}=\left(1-\frac{1}{n_i}\right)Y_{ij}-\frac{1}{n_i}\sum_{j\neq i}Y_{ij}$. Potom náhodná veličina H_{ij} je lineární kombinací nezávislých náhodných veličin s normálním rozdělením a má tedy sama normální rozdělení se střední hodnotou

$$\mathrm{E}H_{ij}=\left(1-\frac{1}{n_i}\right)\mu_i-\frac{n_i-1}{n_i}\mu_i=0$$

a rozptylem

$$\mathrm{var}(H_{ij})=\left(1-\frac{1}{n_i}\right)^2\sigma_i^2+\frac{n_i-1}{n_i^2}\sigma_i^2=\left(1-\frac{1}{n_i}\right)\sigma_i^2.$$

Pak $Z_{ij}=|H_{ij}|$ má polo-normální rozdělení s parametrem $\mathrm{var}(H_{ij})$, což jsme chtěli dokázat.

Leveneův test modifikuje nejprve získané hodnoty následujícím způsobem:

$$|Y_{11} - y_{1.}|, \dots ,|Y_{1n_1} - y_{1.}|$$
$$\vdots$$
$$|Y_{k1} - y_{k.}|, \dots ,|Y_{kn_k} - y_{k.}|,$$

tedy na i-tém řádku odečte od všech hodnot jejich výběrový průměr $y_{i.}$ a poté vezme absolutní hodnoty. Na tento soubor dat se aplikuje ANOVA, která testuje rovnost středních hodnot. V tomto případě jsou střední hodnoty veličin na i-tém řádku (tedy veličin $Z_{ij} = |Y_{ij} - y_{i.}|$) rovny dle Vět 1 a 2

$$\mathrm{E}Z_{ij} = \sqrt{\frac{2}{\pi}\mathrm{var}(H_{ij})} = \sqrt{\frac{2}{\pi}\left(1 - \frac{1}{n_i}\right)}\,\sigma_i.$$

Tedy ve skutečnosti testujeme hypotézu

$$H_0 : \sqrt{\frac{2}{\pi}\left(1 - \frac{1}{n_1}\right)}\,\sigma_1 = \ldots = \sqrt{\frac{2}{\pi}\left(1 - \frac{1}{n_k}\right)}\,\sigma_k$$

proti alternativě, že tomu tak není. Pokud by rozsahy výběrů byly stejné, pak je tato hypotéza ekvivalentní s původní hypotézou

$$H_0 : \sigma_1^2 = \ldots = \sigma_k^2.$$

Nicméně, pro dostatečně velké četnosti n_i se koeficienty $\sqrt{1 - \frac{1}{n_i}}$ nebudou příliš lišit od jedné.

Vlastní test pak probíhá tak, že porovnáme hodnotu testové statistiky

$$W = \frac{(N-k)\sum_{i=1}^{k} n_i(z_{i.} - z_{..})^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Z_{ij} - z_{i.})^2},$$

kde $Z_{ij} = |Y_{ij} - y_{i.}|$, s kritickou hodnotou Fisherova-Snedecorova rozdělení o $k-1$ a $N-k$ stupních volnosti. Nulovou hypotézu potom zamítáme na hladině α, je-li $W \geq F_{k-1,N-k}(\alpha)$, kde $F_{k-1,N-k}(\alpha)$ je příslušná kritická hodnota.

2.3 Předpoklady analýzy rozptylu

Prvním předpokladem analýzy rozptylu je, že máme k nezávislých náhodných výběrů. Jak již bylo řečeno, veličiny $Z_{ij} = |Y_{ij} - y_{i.}|$ a $Z_{ik} = |Y_{ik} - y_{i.}|$ pro $j,k \in \{1, \dots ,n_i\}, j \neq k$ nezávislé nejsou. Lze však ukázat, že korelace mezi nimi je řádu $O(1/n_i^2)$ (viz Fisher, 1920). Ani nulová korelace nám samozřejmě nezaručuje nezávislost, nicméně dle (Levene, 1960) nebude mít korelace tohoto řádu pravděpodobně velký vliv na rozdělení testové statistiky.

Dále požadujeme výběry z normálního rozdělení, což absolutní hodnoty nemohou mít. Dle (Miller, 1997, str. 80) nemá však porušení normality velký vliv. Lze ukázat, že $(k-1)F$, kde F je definováno vzorcem 1.1, má asymptoticky χ^2 rozdělení s $k-1$ stupni volnosti, pokud $n_i \to \infty, i = 1 \dots ,k$, pro jakékoli výchozí rozdělení s konečným rozptylem. Navíc pro dostatečně velká n (a malá k)

jsou kritické hodnoty[1] Fisherova-Snedecorova rozdělení o $k-1$ a $n-k$ stupních volnosti násobené koeficientem $k-1$ větší než příslušné kritické hodnoty χ^2 rozdělení o $k-1$ stupních volnosti. Tedy nezvyšujeme pravděpodobnost zamítnutí nulové hypotézy, respektive nezvyšujeme pravděpodobnost chyby prvního druhu. Z toho plyne, že analýza rozptylu je robustní statistická metoda, což je důvodem robustnosti vlastního Leveneova testu. O tom se přesvědčíme v Kapitole 4.

Předpokládáme i shodné rozptyly. Víme, že rozptyl veličin $Z_{ij} = |Y_{ij} - y_{i.}|$ je dle Vět 1 a 2 roven

$$\mathrm{var}(Z_{ij}) = \left(1 - \frac{2}{\pi}\right)\left(1 - \frac{1}{n_i}\right)\sigma_i^2, i = 1, \ldots, k, j = 1, \ldots, n_i.$$

Tedy tento předpoklad je za platnosti nulové hypotézy splněn, pokud jsou četnosti výběrů stejné. Jinak se lze spoléhat, že pro dostatečně velká n_i se koeficienty $\left(1 - \frac{1}{n_i}\right)$ příliš neliší.

2.4 Další varianty Leveneova testu

Kromě původního Leveneova testu je možné využít některou z jeho variant. Následující dvě, navrženy v (Brown a Forsythe, 1974), jsou založeny na záměně výběrového průměru za jiný odhad, který je v případě symetrického rozdělení odhadem střední hodnoty.

Označme $\tilde{y}_i$ medián i-tého výběru a $\hat{y}_i$ jeho 0,1-useknutý průměr.[2] Definujeme testovou statistiku

$$W = \frac{(N-k)\sum_{i=1}^{k} n_i (z_{i.} - z_{..})^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Z_{ij} - z_{i.})^2}, \tag{2.1}$$

kde Z_{ij} definujeme jako:

- $Z_{ij} = |Y_{ij} - y_{i.}|$, tedy původní Leveneův test, označíme jej W_0;
- $Z_{ij} = |Y_{ij} - \tilde{y}_i|$, tedy odečítáme medián, test označíme W_{50};
- $Z_{ij} = |Y_{ij} - \hat{y}_i|$, tedy odečítáme 0,1-useknutý průměr, test označíme W_{10}.

Dále navrhujeme vynásobení veličin $Z_{ij} = |Y_{ij} - y_{i.}|$ koeficientem $\left(1 - \frac{1}{n_i}\right)^{-\frac{1}{2}}$. Díky tomu budou mít tyto veličiny střední hodnotu $\sqrt{\frac{2}{\pi}}\,\sigma_i$ a tedy testovaná hypotéza rovnosti středních hodnot bude ekvivalentní hypotéze o shodnosti rozptylů i při nevyvážených výběrech. Navíc rozptyly veličin na i-tém řádku budou rovny $\left(1 - \frac{2}{\pi}\right)\sigma_i^2$, a tedy shodné právě tehdy, když původní rozptyly byly shodné. Tato modifikace nijak nemění testovou statistiku, pokud jsou výběry vyvážené, a zdá se, že by mohla být lepší, máme-li silně nevyvážené výběry. Samozřejmě i tato modifikace může být provedena ve všech třech variantách, tedy lze odečítat výběrový průměr, výběrový medián nebo výběrový 0,1-useknutý průměr.

[1] Připomínáme, že kritická hodnota je $1-\alpha$ kvantilem daného rozdělení.

[2] Obecně alfa-useknutý průměr z n hodnot je vypočítán tak, že nejprve odebereme $\lfloor \alpha \cdot n \rfloor$ nejmenších a stejný počet největších pozorování, a ze zbylých hodnot vypočítáme průměr. Obyčejný průměr lze pak považovat za 0-useknutý průměr a medián za 0,5-useknutý průměr.

Mějme stejnou testovou statistiku 2.1, kde

- $Z_{ij} = \left(1 - \frac{1}{n_i}\right)^{-\frac{1}{2}} |Y_{ij} - y_{i.}|$, test označíme $\overline{W}_0$;
- $Z_{ij} = \left(1 - \frac{1}{n_i}\right)^{-\frac{1}{2}} |Y_{ij} - \tilde{y_{i.}}|$, test označíme $\overline{W}_{50}$.

V Kapitole 4 pomocí simulací zjistíme, jak se tyto varianty Leveneova testu chovají. Navíc je porovnáme s dalšími testy shodnosti rozptylů, které budou popsány v následující kapitole.

Kapitola 3

Další testy shodnosti rozptylů

V této kapitole, jejíž hlavní část je převzata z (Anděl, 2003), popíšeme další tři možné testy shodnosti rozptylů. Všechny předpokládají nezávislé náhodné výběry z normálního rozdělení.

3.1 Bartlettův test

Předpokládejme, že máme k nezávislých náhodných výběrů z normálního rozdělení. Označíme

$$s_i^2 = \frac{1}{n_i - 1}\left(\sum_{j=1}^{n_i} Y_{ij}^2 - n_i y_{i.}^2\right), \quad i = 1, \dots, k,$$

$$s^2 = \frac{1}{N-k}\sum_{i=1}^{k}(n_i - 1)s_i^2,$$

$$C = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{N-k}\right).$$

Pokud $n_i > 6$ pro každé i, má veličina

$$B = \frac{1}{C}\left[(N-k)\ln s^2 - \sum_{i=1}^{k}(n_i - 1)\ln s_i^2\right]$$

za platnosti nulové hypotézy shodnosti rozptylů přibližně χ^2_{k-1} rozdělení. Tedy že pro četnosti výběrů rostoucí nade všechny meze konverguje distribuční funkce veličiny B k distribuční funkci veličiny, která má χ^2_{k-1} rozdělení.[1]

Bartlettův test je narozdíl od Leveneova testu velmi citlivý na porušení normality. To znamená, že máme-li náhodné výběry, které nepocházejí z normálního rozdělení, pak Bartlettův test nemusí dodržovat (ani asymptoticky) požadovanou hladinu α, o čemž se přesvědčíme v Kapitole 4.

[1] To jest $B \xrightarrow{d} \chi^2_{k-1}$, tedy B konverguje v distribuci k χ^2 rozdělení o $k-1$ stupních volnosti.

3.2 Hartleyův a Cochranův test

Mějme stejné předpoklady a značení, jako v 3.1. Nechť jsou navíc všechny četnosti n_i stejné a označme $m = n_1 - 1$.

V Hartleyově testu porovnáme hodnotu testové statistiky

$$F_{max} = \frac{\max_{j=1,\dots,k} s_j^2}{\min_{i=1,\dots,k} s_i^2}$$

s kritickými hodnotami $h_{k,m}(\alpha)$, které jsou uvedeny například v (Anděl, 2003, Tabulka T13).

Cochranův test je založen na statistice

$$G_{max} = \frac{\max_{j=1,\dots,k} s_j^2}{\sum_{i=1}^{k} s_i^2}.$$

Kritické hodnoty $c_{k,m}(\alpha)$ pro tuto statistiku jsou uvedeny například v (Anděl, 2003, Tabulka T14).

Kapitola 4

Simulační experimenty

Návrh simulačních experimentů je inspirován simulacemi uvedenými v (Brown a Forsythe, 1974), kde jsou porovnány některé modifikace Leveneova testu s testy shodnosti rozptylů, které navrhl M. W. J. Layard v (Layard, 1973), a s F-testem. My však porovnáváme Leveneův test (v modifikacích uvedených v této práci) s Bartlettovým, Hartleyovým a Cochranovým testem.

4.1 Návrh simulačních experimentů

Pomocí programu R (R Core Team, 2013, verze 2.15.3) byla vygenerována pseudonáhodná čísla z normálního rozdělení, Studentova $\mathbf{t}_4$ rozdělení a χ^2_4 rozdělení. V tomto programu byly provedeny také veškeré testy a výpočty.

Z každého rozdělní byly simulovány výběry s četnostmi (10,10), (20,20), (40,40), (1000,1000), (10,20), (10,40), (10,50), (20,40), (200,600) a to pro poměry rozptylů 1:1, 1:2, 1:4 pro všechny a navíc 2:1, 4:1 pro nevyvážené soubory. V případech χ^2_4 a $\mathbf{t}_4$ rozdělení byl druhý soubor čísel vynásoben vhodným koeficientem tak, aby bylo dosaženo požadovaného poměru rozptylů. V případě normálního rozdělení byl výběr generován přímo s požadovaným rozptylem.

Pro všechny kombinace bylo generováno 1000 dvojic výběrů a za platnosti nulové hypotézy pak 10 000. Jednotlivé testy byly prováděny na stejných výběrech.

Simulacemi byly porovnány následující testové statistiky:

1. Leveneův test založený na výběrovém průměru - označený W_0,
2. Leveneův test založený na výběrovém mediánu - označený W_{50},
3. Leveneův test založený na 0,1-useknutém průměru - označený W_{10},
4. Leveneův test s vynásobením a centrováním pomocí průměru (pouze nevyvážené výběry) - označený $\overline{W}_0$,
5. Leveneův test s vynásobením a centrováním pomocí mediánu (pouze nevyvážené výběry) - označený $\overline{W}_{50}$,
6. Bartlettův test - označený Bartlett,
7. Hartleyův test (pouze vyvážené výběry) - označený Hartley,
8. Cochranův test (pouze vyvážené výběry).

Netestovali jsme F-test popsaný v úvodu této práce. Jedním z důvodů bylo, že jej lze použít pouze pro dva náhodné výběry. Dalším důvodem byla skutečnost, že byl testován v (Brown a Forsythe, 1974), kde je uvedeno, že dával v podstatě stejné výsledky jako Bartlettův test.

4.2 Výsledky simulací

Výsledky jsou shrnuty v tabulkách 4.1–4.6, kde jsou uvedeny počty zamítnutí nulové hypotézy v procentech. Vzhledem k faktu, že testy byly prováděny na hladině $\alpha = 0{,}05$, nejsou uvedeny výsledky testů, které překročily hladinu 0,08 - podobně jako v (Brown a Forsythe, 1974). Dále jsou v závorkách testy pohybující se na hladině $0{,}07 - 0{,}08$.

Testy na hladině 0,01 vycházely analogicky jako testy na hladině 0,05, proto nejsou uvedeny vůbec. Také není uveden Cochranův test, neboť zamítl téměř vždy stejný počet výběrů, jako Bartlettův test (lišily se vždy nejvýše u jednoho výběru).

Tabulky pro vyvážené a nevyvážené výběry jsou odděleny, protože Hartleyův test je možné použít pouze pro vyvážené výběry, naopak poslední dvě modifikace Leveneova testu $\overline{W}_0$ a $\overline{W}_{50}$ dávají pro vyvážené výběry samozřejmě stejné výsledky, jako W_0, respektive W_{50}.

Komentáře získaných výsledků

Z výsledků je patrné, že jediné testy, které při všech simulacích držely požadovanou hladinu 0,05, jsou Leveneův test s centrováním pomocí mediánu W_{50}, případně jeho modifikace s vynásobením $\overline{W}_{50}$. Velmi dobře se choval i Leveneův test založený na 0,1-useknutém průměru W_{10}, nicméně ne vždy udržel požadovanou hladinu viz tabulky 4.5 a 4.6. Podívejme se nyní podrobněji na výsledky testů vzhledem k rozdělení výběrů.

V případě normálního rozdělení (tabulky 4.1 a 4.2) hladinu drží všechny testy, ale síla testů se mírně liší, nejméně silnými testy jsou W_{50} a $\overline{W}_{50}$. Pokud jsme si tedy jisti normalitou dat, pak volíme nejspíše původní Leveneův test, Bartlettův test nebo Hartleyův test. Pro $k = 2$ lze také použít F-test popsaný v úvodu této práce, ten je přesný.

Máme-li výběry pocházející ze Studentova **t** rozdělení (tabulky 4.3 a 4.4), pak hladinu drží v podstatě (pro dostatečně velké četnosti) všechny modifikace Leveneova testu. Naproti tomu Bartlettův test ji nedrží vůbec, tedy naše simulace potvrzují jeho citlivost na porušení normality dat. Ze stejného důvodu je nepoužitelný i Hartleyův test.

Normalita byla nejvíce porušena, pokud výběry pocházely z χ^2 rozdělení (tabulky 4.5 a 4.6), které je narozdíl od předchozích dvou rozdělení asymetrické. Nicméně i tady testy W_{50} a $\overline{W}_{50}$ držely zvolenou hladinu, pro dostatečně velké četnosti ji držel i test W_{10}. Naopak hladinu nedržel původní Leveneův test W_0 (ani jeho modifikace $\overline{W}_0$), Bartlettův a Hartleyův test.

Porovnání testů W_0 a $\overline{W}_0$

Zajímavým výsledkem může být i porovnání testů W_0 s $\overline{W}_0$, respektive W_{50} s $\overline{W}_{50}$. Chování těchto dvojic je analogické. Porovnejme například testy W_0 a $\overline{W}_0$. Oba drží v podstatě stejnou hladinu, ale jejich síla je různá. Ukazuje se, že test W_0 je silnější, pochází-li náhodný výběr s menší četností z rozdělení s menším rozptylem. Tento výsledek lze snad zdůvodnit tím, že pomocí W_0 testujeme hypotézu

$$H_0 : \sqrt{\frac{2}{\pi}\left(1 - \frac{1}{n_1}\right)}\,\sigma_1 = \sqrt{\frac{2}{\pi}\left(1 - \frac{1}{n_2}\right)}\,\sigma_2,$$

zatímco $\overline{W}_0$ testuje hypotézu

$$H_0 : \sqrt{\frac{2}{\pi}}\,\sigma_1 = \sqrt{\frac{2}{\pi}}\,\sigma_2.$$

Máme-li tedy $2 \leq n_1 < n_2$ a $0 < \sigma_1^2 < \sigma_2^2$, potom

$$1 - \frac{1}{n_1} < 1 - \frac{1}{n_2} \Rightarrow 1 < \frac{\sqrt{1 - \frac{1}{n_2}}}{\sqrt{1 - \frac{1}{n_1}}}.$$

A tedy

$$\sigma_2 \frac{\sqrt{1 - \frac{1}{n_2}}}{\sqrt{1 - \frac{1}{n_1}}} - \sigma_1 > \sigma_2 - \sigma_1 > 0,$$

z čehož plyne, že test W_0 detekuje větší rozdíl.

Naopak test $\overline{W}_0$ je silnější, pochází-li výběr s menší četností z rozdělení s větším rozptylem. Pokusme se zdůvodnit i tuto skutečnost, přestože je o něco složitější. Předpokládejme, že $2 \leq n_1 < n_2$ a $\sigma_1 > \sigma_2 > 0$. Platí opět

$$\sigma_2 \frac{\sqrt{1 - \frac{1}{n_2}}}{\sqrt{1 - \frac{1}{n_1}}} - \sigma_1 > \sigma_2 - \sigma_1,$$

ale $\sigma_2 - \sigma_1 < 0$, a tedy nevíme skoro nic. Kdyby byla levá strana nerovnosti menší než 0, pak by bylo vše jasné. To však vždy platit nemusí. Ve skutečnosti by stačilo, kdyby platilo

$$\left| \sigma_2 \frac{\sqrt{1 - \frac{1}{n_2}}}{\sqrt{1 - \frac{1}{n_1}}} - \sigma_1 \right| < |\sigma_2 - \sigma_1| = \sigma_1 - \sigma_2,$$

pak by test $\overline{W}_0$ detekoval větší rozdíl, a proto by mohl být silnější. K tomu potřebujeme ještě vědět, zda platí

$$\sigma_1 - \sigma_2 > \sigma_2 \frac{\sqrt{1 - \frac{1}{n_2}}}{\sqrt{1 - \frac{1}{n_1}}} - \sigma_1,$$

což je ekvivalentní s

$$2\sigma_1 > \sigma_2 \left(1 + \frac{\sqrt{1-\frac{1}{n_2}}}{\sqrt{1-\frac{1}{n_1}}}\right),$$

respektive s

$$\frac{\sigma_1}{\sigma_2} > \frac{1}{2}\left(1 + \frac{\sqrt{1-\frac{1}{n_2}}}{\sqrt{1-\frac{1}{n_1}}}\right).$$

Pravá strana poslední uvedené nerovnosti bude největší, pokud n_1 bude nejmenší možné a n_2 největší možné. V našich simulacích bylo n_1 nejméně 10 a n_2 nejvíce 200. Položíme $n_1 = 10$ a n_2 necháme jít k nekonečnu, tedy $n_2 \to \infty$. Potom by mělo platit

$$\frac{\sigma_1}{\sigma_2} > \frac{1}{2}\left(1 + \frac{1}{\sqrt{0{,}9}}\right) \doteq 1{,}02705.$$

A tedy stačí, aby $\frac{\sigma_1^2}{\sigma_2^2} > 1{,}055$.

V našich simulacích bylo vždy $\frac{\sigma_1^2}{\sigma_2^2} \geq 2$, proto všechny naše výsledky ukazují, že test $\overline{W}_0$ je silnější než test W_0, je-li $n_1 < n_2$ a $\sigma_1^2 > \sigma_2^2$. Z uvedeného poměru, který je navíc důsledkem velmi extrémní situace, je rovněž vidět, že pokud daná nerovnost není splněna, jsou si rozptyly stejně velmi blízké.

Poznamenejme, že nejhorší situace bychom dosáhli, kdybychom položili $n_1 = 2$ a $n_2 \to \infty$. Daný poměr by potom vycházel $\frac{1}{2}(1+\sqrt{2}) \doteq 1{,}21$, respektive jeho druhá mocnina by byla přibližně 1,46. Ale v praxi nelze očekávat, že bychom zkoumali rozptyl, máme-li pouze dvě hodnoty.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	Bartlett	Hartley
10:10	1:1	5,9	3,4	5,3	4,9	6,1
	1:2	15,4	10,1	14,6	16,2	18,8
	1:4	40,5	31,0	39,0	50,7	55,2
20:20	1:1	5,7	4,5	5,5	5,1	5,7
	1:2	27,2	21,8	25,8	30,6	33,2
	1:4	77,1	71,6	75,9	85,0	85,9
40:40	1:1	5,2	4,3	5,1	5,3	5,6
	1:2	50,4	46,6	49,0	58,0	58,7
	1:4	98,0	97,4	97,8	99,1	99,2
1000:1000	1:1	4,8	4,8	4,8	5,1	5,1
	1:2	100,0	100,0	100,0	100,0	100,0

Tabulka 4.1: Počet zamítnutí nulové hypotézy (v %) pro normální rozdělení, vyvážené výběry a hladinu $\alpha = 0{,}05$.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	$\overline{W}_0$	$\overline{W}_{50}$	Bartlett
10:20	1:1	5,3	3,7	4,9	5,4	3,8	5,1
	1:2	15,1	10,5	13,8	13,1	9,4	16,9
	1:4	48,8	40,5	48,2	45,3	37,7	58,7
	2:1	21,9	15,7	20,5	24,7	18,7	23,0
	4:1	63,4	52,6	61,2	66,1	55,8	67,6
10:40	1:1	5,5	4,5	5,3	5,9	4,7	5,2
	1:2	17,4	16,6	17,1	14,8	13,1	21,7
	1:4	59,6	55,2	58,0	53,4	50,1	68,4
	2:1	26,2	20,8	25,2	31,5	25,8	26,9
	4:1	71,4	65,6	70,0	75,4	70,0	75,5
10:50	1:1	5,6	4,6	5,4	6,1	5,0	5,3
	1:2	17,7	17,0	17,9	14,5	13,6	22,0
	1:4	61,4	58,4	61,1	55,1	52,6	72,2
	2:1	27,1	23,0	25,7	32,8	27,6	26,7
	4:1	73,6	69,0	72,4	77,1	72,9	76,5
20:40	1:1	5,5	4,3	5,3	5,6	4,4	5,2
	1:2	33,5	31,1	33,5	32,1	29,9	39,2
	1:4	87,7	86,0	87,3	87,0	84,7	93,5
	2:1	38,2	32,5	37,2	39,9	35,0	38,6
	4:1	88,0	86,2	87,4	88,8	86,9	92,1
200:600	1:1	5,4	5,3	5,4	5,3	5,3	5,3
	1:2	100,0	100,0	100,0	100,0	100,0	100,0
	2:1	100,0	100,0	100,0	100,0	100,0	99,9

Tabulka 4.2: Počet zamítnutí nulové hypotézy (v %) pro normální rozdělení, nevyvážené výběry a hladinu $\alpha = 0{,}05$.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	Bartlett	Hartley
10:10	1:1	6,6	3,7	5,5	16,0	18,5
	1:2	13,3	9,2	12,2		
	1:4	32,2	23,3	29,0		
20:20	1:1	6,0	4,1	5,0	21,3	22,4
	1:2	20,1	16,4	18,8		
	1:4	59,3	52,1	56,2		
40:40	1:1	5,4	4,3	4,9	24,2	38,6
	1:2	35,7	32,7	33,6		
	1:4	86,2	85,1	85,9		
1000:1000	1:1	5,4	5,3	5,3	38,6	38,6
	1:2	100,0	100,0	100,0		

Tabulka 4.3: Počet zamítnutí nulové hypotézy (v %) pro Studentovo **t** rozdělení o 4 stupních volnosti, vyvážené výběry a hladinu $\alpha = 0{,}05$.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	$\overline{W}_0$	$\overline{W}_{50}$	Bartlett
10:20	1:1	6,1	3,6	4,9	6,5	3,9	17,5
	1:2	12,5	9,4	11,3	11,2	8,6	
	1:4	30,9	26,5	28,8	29,3	24,4	
	2:1	21,5	14,6	18,9	23,3	17,0	
	4:1	52,5	43,1	48,7	54,5	45,0	
10:40	1:1	5,7	3,7	4,5	6,4	4,2	18,1
	1:2	9,6	8,4	9,4	8,6	6,7	
	1:4	34,5	30,7	33,1	30,8	27,8	
	2:1	22,7	17,9	20,4	26,5	21,1	
	4:1	58,6	53,3	56,9	63,1	56,6	
10:50	1:1	5,9	3,9	4,7	6,6	4,5	18,4
	1:2	8,6	7,7	8,2	7,3	6,1	
	1:4	32,5	30,0	31,0	28,6	26,6	
	2:1	21,8	17,1	19,7	24,7	20,5	
	4:1	57,7	51,9	55,7	62,8	56,6	
20:40	1:1	6,0	4,6	5,2	6,0	4,5	22,3
	1:2	22,5	20,6	21,2	21,0	19,1	
	1:4	65,8	62,8	64,7	63,6	61,3	
	2:1	28,8	24,2	26,4	30,3	25,9	
	4:1	73,3	69,0	71,3	75,4	70,6	
200:600	1:1	5,3	5,1	5,2	5,3	5,1	34,6
	1:2	99,2	99,3	99,3	99,2	99,3	
	1:4	100,0	100,0	100,0	100,0	100,0	
	2:1	98,5	98,5	98,5	98,6	98,6	
	4:1	100,0	100,0	100,0	100,0	100,0	

Tabulka 4.4: Počet zamítnutí nulové hypotézy (v %) pro Studentovo **t** rozdělení o 4 stupních volnosti, nevyvážené výběry a hladinu $\alpha = 0{,}05$.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	Bartlett	Hartley
10:10	1:1	10,8	4,4	8,0	14,6	17,0
	1:2		10,8			
	1:4		25,4			
20:20	1:1	10,6	4,7	7,6	17,5	18,8
	1:2		17,2	(22,8)		
	1:4		51,9	(59,4)		
40:40	1:1	9,8	4,6	6,9	19,3	19,8
	1:2		34,2	40,0		
	1:4		87,2	90,0		
1000:1000	1:1	9,9	4,9	6,6	21,7	21,7
	1:2		100,0	100,0		
	1:42		100,0	100,0		

Tabulka 4.5: Počet zamítnutí nulové hypotézy (v %) pro χ^2 rozdělení o 4 stupních volnosti, vyvážené výběry a hladinu $\alpha = 0{,}05$.

n_1,n_2	$\sigma_1^2 : \sigma_2^2$	W_0	W_{50}	W_{10}	$\overline{W}_0$	$\overline{W}_{50}$	Bartlett
10:20	1:1	10,5	4,3	7,7	10,6	4,4	15,7
	1:2		12,1	(17,3)		10,6	
	1:4		31,4	(40,1)		28,9	
	2:1		15,3	(21,0)		16,8	
	4:1		40,7	(49,2)		42,7	
10:40	1:1	9,6	4,2	6,6	9,8	4,6	15,7
	1:2		9,5	14,1		8,3	
	1:4		33,0	41,0		29,2	
	2:1		16,7	21,2		19,4	
	4:1		52,7	59,3		57,2	
10:50	1:1	9,4	4,3	6,7	10,4	4,7	15,8
	1:2		12,1	15,2		10,6	
	1:4		34,5	42,9		31,5	
	2:1		18,5	23,2		21,9	
	4:1		50,9	57,0		55,9	
20:40	1:1	10,1	4,4	7,0	10,3	4,5	17,8
	1:2		22,1	(25,8)		20,9	
	1:4		64,1	(71,4)		62,3	
	2:1		22,6	(29,0)		23,6	
	4:1		71,1	(76,3)		72,2	
200:600	1:1	9,2	4,5	6,1	9,2	4,6	20,1
	1:2		99,3	99,3		99,1	
	1:4		100,0	100,0		100,0	
	2:1		99,6	99,7		99,7	
	4:1		100,0	100,0		100,0	

Tabulka 4.6: Počet zamítnutí nulové hypotézy (v %) pro χ^2 rozdělení o 4 stupních volnosti, nevyvážené výběry a hladinu $\alpha = 0{,}05$.

Závěr

Cílem této práce bylo popsat a vysvětlit Leveneův test a některé jeho možné modifikace. Pomocí simulačních experimentů pak ukázat, zda jednotlivé testy drží předepsanou hladinu a porovnat je mezi sebou z hlediska síly testu.

Uvedli jsme původní Leveneův test a popsali jeho princip. Dále jsme předložili modifikace navržené v (Brown a Forsythe, 1974), navrhli další možnou modifikaci na základě výpočtů středních hodnot a rozptylů veličin používaných v původním testu.

Díky simulacím jsme ukázali, že dvě z modifikací Leveneova testu drží zvolenou hladinu, máme-li náhodné výběry nejen z normálního, ale dokonce i ze Studentova $\mathbf{t}$ a z χ^2 rozdělení (obě o 4 stupních volnosti). To může být jedním z důvodů, proč je Leveneův test (v modifikaci se záměnou výběrového průměru za medián) doporučován k testování hypotézy k nezávislých náhodných výběrů i v situaci, kdy není zaručena normalita dat.

Seznam použité literatury

Anděl, J. (2003). *Statistické metody.* Vydání třetí. Matfyzpress, Praha. ISBN 80-85863-27-8.

Anděl, J. (2007). *Základy matematické statistiky.* Druhé opravené vydání. Matfyzpress, Praha. ISBN 80-7378-001-1.

Brown, M. B. a Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, **69**, 364–367.

Fisher, R. A. (1920). A Mathematical Examination of the Methods of determining the Accuracy of an Observation by the Mean Error, and by the Mean Square Error. *Monthly Notices of the Royal Astronomical Society*, **80**, 758–770.

Layard, M. W. J. (1973). Robust Large-Sample Tests for Homogeneity of Variances. *Journal of the American Statistical Association*, **68**, 195–198.

Levene, H. (1960). Robust Tests for Equality of Variances. In I. Olkin, Palo Alto, C., editor, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling.* Stanford University Press, 278-292.

Miller, R. (1997). *Beyond ANOVA : Basics of Applied Statistics.* Chapman & Hall, London New York. ISBN 0412070111.

R Core Team (2013). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`. ISBN 3-900051-07-0.

Seznam tabulek

4.1 Počet zamítnutí nulové hypotézy (v %) pro normální rozdělení, vyvážené výběry a hladinu $\alpha = 0{,}05$. 15
4.2 Počet zamítnutí nulové hypotézy (v %) pro normální rozdělení, nevyvážené výběry a hladinu $\alpha = 0{,}05$. 15
4.3 Počet zamítnutí nulové hypotézy (v %) pro Studentovo **t** rozdělení o 4 stupních volnosti, vyvážené výběry a hladinu $\alpha = 0{,}05$. 16
4.4 Počet zamítnutí nulové hypotézy (v %) pro Studentovo **t** rozdělení o 4 stupních volnosti, nevyvážené výběry a hladinu $\alpha = 0{,}05$. . . . 16
4.5 Počet zamítnutí nulové hypotézy (v %) pro χ^2 rozdělení o 4 stupních volnosti, vyvážené výběry a hladinu $\alpha = 0{,}05$. 17
4.6 Počet zamítnutí nulové hypotézy (v %) pro χ^2 rozdělení o 4 stupních volnosti, nevyvážené výběry a hladinu $\alpha = 0{,}05$. 17